

Remote Sensing Data Visualization, Fusion and Analysis via Giovanni

G. Leptoukh^a, V. Zubko^b, A. Gopalan^b, M. Khayat^b

NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA – Gregory.Leptoukh@nasa.gov
RS Information Systems, Inc – (vzubko, agopalan, mkhayat@pop600.gsfc.nasa.gov)

Abstract –We describe Giovanni, the NASA Goddard developed online visualization and analysis tool that allows users explore various phenomena without learning remote sensing data formats and downloading voluminous data. Using MODIS aerosol data as an example, we formulate an approach to the data fusion for Giovanni to further enrich online multi-sensor remote sensing data comparison and analysis.

Keywords: Remote sensing, Data distribution, Data tools, Data fusion, Value-added services, Aerosol optical thickness, MODIS, Terra, Aqua, Giovanni

1. INTRODUCTION

The NASA Earth Observing System (EOS) multi-satellite data archives are indispensable for studying regional or global atmospheric phenomena. Until recently, using this data required being able to locate and retrieve the relevant data coupled with a detailed understanding of the data's complicated internal structure. Consequently, this data was largely unusable to the public at large as gaining the knowledge required to carry out the data reduction is a time-consuming task which must be undertaken well in advance. Even for experienced users analysis of multi-sensor data sets that are typically in different formats, structures, and resolutions is a daunting task.

The NASA Goddard Earth Sciences Data and Information Services Center (GES DISC) has recognized this complexity and has taken a major step towards developing a user friendly Web interface that allows users to perform interactive analysis online without downloading any data, or needing to understand complicated data structures. The Goddard Interactive Online Visualization and Analysis Infrastructure or "Giovanni" addresses these objectives (Acker and Leptoukh, 2007). Giovanni (<http://giovanni.gsfc.nasa.gov>) has successfully demonstrated its utility as an interactive, online, analysis tool for data users to facilitate a wide spectrum of users in research, education, and the curious internet surfer.

One of the expressed interests of users worldwide has been to combine and fuse data from multiple sensors using Giovanni. GES DISC as a significant data archive location is uniquely positioned to address this need using data fusion (DF) techniques. DF is the intelligent merging or integration of data from multiple sources to extract more or better information than would be possible from the individual sources. With the vast quantity of satellite data sets available from numerous missions and sensors, many of which are complementary to each other, there is an increasingly critical need to combine to derive the most optimal benefits from the data. Often, information provided by an individual sensor may be incomplete, inconsistent, inadequate,

and/or imprecise. Fusing of multi-sensor data, e.g., Aerosol Optical Thickness (AOT), can potentially create more consistent, reliable, and complete picture of the space-time evolution of the underlying geophysical process. Missing data from one sensor could be intelligently "filled in" with available co-located data from another sensor to produce a better estimate of geophysical parameters.

The work described in this paper is part of the larger effort to enable DF in Giovanni. We provide a quick overview of Giovanni capabilities with emphasis on our plans for the added DF capability, using data from the Moderate Resolution Imaging Spectroradiometer (MODIS) Terra and Aqua gridded daily mean AOT products. Our objective here is on increasing the spatial coverage: filling orbital and other gaps through DF and incorporating the simplest DF algorithm of the Terra and Aqua AOTs.

2. GIOVANNI

Giovanni offers a user access, visualization, and analysis tools with in a user friendly GUI. With a few mouse clicks, the user can easily obtain various remote sensing or model information from around the globe. Users become explorers of data interactively and online without the overhead of first having to download the data. Giovanni also eliminates the need to understand complicated data formats before one is able to initiate the intended analysis. Access is provided through a common Web browser, so the user does not need special applications beyond what is available on a typical personal computer.

From a web page the user is able to select the spatial area "box" for the desired region via a Java image map applet or manually enter the coordinates defining the bounding box. The user also selects the temporal range for the data, one or more parameters from this data set, and the desired output type (ASCII or one of several plot types). For the plot selections, several color options are also available. The user is then able to refine this analysis and download the results. ASCII output is useful for GIS or other user applications, and the plots generated can be extracted into the user's final report or paper. For more detailed analysis, links to the data are available so the user can download the entire data set for further local analysis. Depending on the choice of parameters, the majority of users will see the online results in a matter of seconds while online manipulation of larger amounts of data (either spatially or temporally) may take several minutes. Even in this more extreme case, the time from the inception of an analysis idea to actually seeing the results is drastically reduced and the most tedious aspects of the analysis are issues that the user bypasses in its entirety!

The MODIS Online Visualization and Analysis System (MOVAS) is an operational instance available through Giovanni since September 2003. MOVAS allows scientists and researchers to

easily access, visualize and analyze MODIS Level-3 atmospheric daily and monthly products helping them for example to understand seasonal-to-inter-annual variation of atmospheric parameters ranging from aerosol to clouds. MOVAS can provide information at every single point and in any rectangular area within the data domain, which allows researchers to conduct nearly unlimited investigations.

MOVAS as a Giovanni instance includes advanced features with capabilities for performing intercomparison analyses between parameters extracted from MODIS instruments onboard two different satellites Terra and Aqua, as well as those from the Goddard Chemistry Aerosol Radiation and Transport (GOCART) model. As such, this instance is an excellent candidate for the implementation and fielding of an initial DF capability.

3. DATA FUSION

3.1 Data: MODIS Terra and Aqua AOT

The Level-3, gridded daily mean AOT data used in this study (Terra and Aqua MODIS Collection 005 Daily Global Gridded Products MOD08_D3.005 and MYD08_D3.005) were acquired by MODIS onboard the NASA Terra and Aqua satellites. Both satellites are polar-orbiting, with Terra on a descending orbit (southward) over the equator at about 10:30 local sun time and Aqua on an ascending orbit (northward) over the equator at about 13:30 local sun time. Terra MODIS has been making global aerosol measurements since February 2000 and Aqua MODIS since July 2002. The daily Level-3 data are space-time aggregated from the Level-2 data (nominal resolution of 10 km x 10 km) to a 1° x 1° resolution (Remer et al., 2005). MODIS uses reflectance measurements made in the visible portion of the spectrum to retrieve aerosol information. Thus, MODIS aerosol measurements are available for daytime only. The daily Level-3 data contain statistics derived from the Level-2 atmospheric products. Any Level-2, standard 5-minute data file that overlaps any part of a data day (0000 to 2400 UTC) is included in the statistics (various moments of the AOT distribution, e.g., quality-weighted mean and standard deviation, number of counts), which are computed within 1° x 1° grid boxes on an equal-angle latitude-longitude projection. The standard deviation is a measure of the variability of AOT in that grid box over that time scale. We used it in the current study as a conservative estimate of the AOT variability for the merging of data from Terra and Aqua MODIS.

Giovanni can be used to rapidly and efficiently create and visualize daily global 1° x 1° maps of AOT (at 0.55 micron) using Terra and Aqua MODIS Level-3 data products. Actually, we used Giovanni during the course of this study to identify interesting cases for data fusion. The typical large gaps (especially near the equatorial regions) in the AOT daily mean field for both Terra MODIS and Aqua MODIS result from a combination of factors, including gaps between swaths from different orbits, and problems in AOT retrievals due to sun glint (over water), cloud cover, or very bright surfaces like deserts (Hsu et al., 2004).

This paper focused on two, 20° latitude by 30° longitude regions, each corresponding to a MODIS subset scene. Subsets were used because our fusion algorithm (see next section) used Optimal Interpolation (OI) to fill in the data gaps, and OI involved the inversion of matrices of large dimensions and was, thus,

computationally expensive to apply to the entire global grid. These two scenes (Subsets 1 and 2) were selected for their variations in the spatial gradients of the AOT field and in the fractional size and distribution of gaps. Subset 1 (Atlantic Ocean off the coast of western Africa) contained mostly regular gaps, whereas Subset 2 (west-central Africa) contained mostly irregular gaps.

3.2 Fusion Algorithm

Our approach to data fusion is (1) to merge the data and then (2) interpolate to fill the gaps. This sequence is optimal in the sense that it preserves original data information most (least distortion).

Data Merging. For merging the data sets, we used weighted averaging, which is a family of methods based on arithmetic combinations of input values, such as linear combinations, weighted multiplication or ratios, and maximum likelihood estimate (MLE). The MLE emphasizes the use of different sources of data using statistics such as mean, standard deviation, and number of counts. For isotropic uncertainty, the MLE can provide a good approximation of the actual estimate of a feature from multiple observations. The MLE requires minimal a priori information, and it is easy to incorporate user-supplied weights for the data sources (Chu and Aggarwal, 1993). For a set of N independent observations (F_k) of the same parameter, the MLE estimate is:

$$F_{MLE} = \sum_{k=1}^N \frac{F_k}{\sigma_k^2} \Big/ \sum_{k=1}^N \frac{1}{\sigma_k^2} \quad (1)$$

where σ_k is the variance of the Gaussian noise affecting the observations. The σ_k is computed for each cell selected for fusion, and the expected estimate of F is calculated using Eq. 1.

Spatial Interpolation for Filling Data Gaps. For filling data gaps, we used the method of Optimal Interpolation (OI), which takes into account spatial correlations in the data. The approach was first introduced to obtain a more accurate objective analysis of meteorological and oceanographic data (Bretherton et al., 1976). It is based on the Gauss-Markov theorem, which determines that an unbiased estimate of the field of interest is linear in the data and has the minimum variance, given the expected value and covariance of both the field and the data. The method assumes that the observational data are spatially correlated, i.e., data that are close to each other are highly correlated, and also takes into account the observation-to-background error variance. It also requires a specification of the spatial correlation length. The method is "optimal" because, if the correlation function is an accurate model of the spatial relationship of the data and if the assumption about noise accurately reflects the level of actual noise in the data, then the method yields the least expected error of the linear estimate. Thus, the data estimate at the analysis point is a linear combination of data at the points of availability:

$$F_j^* = \sum_{i=1}^N w_{ji} F_i \quad (2)$$

where F_j^* is the estimate at the analysis point j ($j=1..m$, where m is the number of analysis points), F_i is the data value at point i . Note that the coefficients w_{ji} are found by solving a linear system, which requires inverting the covariance matrix. Thus the OI method can be quite time consuming for large data sets.

3.3 Results

The advantages of merging the daily mean AOT fields from MODIS Terra and Aqua first and then using OI to fill the gaps are illustrated in Figures 1 and 2 for, respectively, Subset 1 (mostly regular gaps) and Subset 2 (mostly irregular gaps). When the original Terra (Figure 1a) or Aqua (Figure 2a) AOT fields were interpolated individually using OI to obtain a complete field, the interpolation errors (Figures 1c or 2c and Figures 1f or 2f) associated with the gap regions were large, especially around the gaps centers (farthest away from the available measurements). The errors decreased towards the gap edges. The advantage of merging two data sets is shown in Figures 1g and 2g by a large reduction in the percentage of pixels with no data. Figures 1h and 2h show the AOT fields resulted from the respective merged ones after applying the OI inside the gaps in the merged fields. Compared with the interpolation errors of individually interpolated Terra or Aqua AOT fields (Figures 1c or 1f and Figures 2c or 2f), the errors of the merged and then interpolated AOT fields (Figures 1i and 2i) were much smaller.

4. CONCLUSIONS

The benefit of first merging data and then interpolating to fill in the remaining gaps was demonstrated to account for the variability of a measured parameter, compared to interpolating in gaps in single datasets and then merging the results. This multi-sensor data fusion approach will be incorporated in Giovanni.

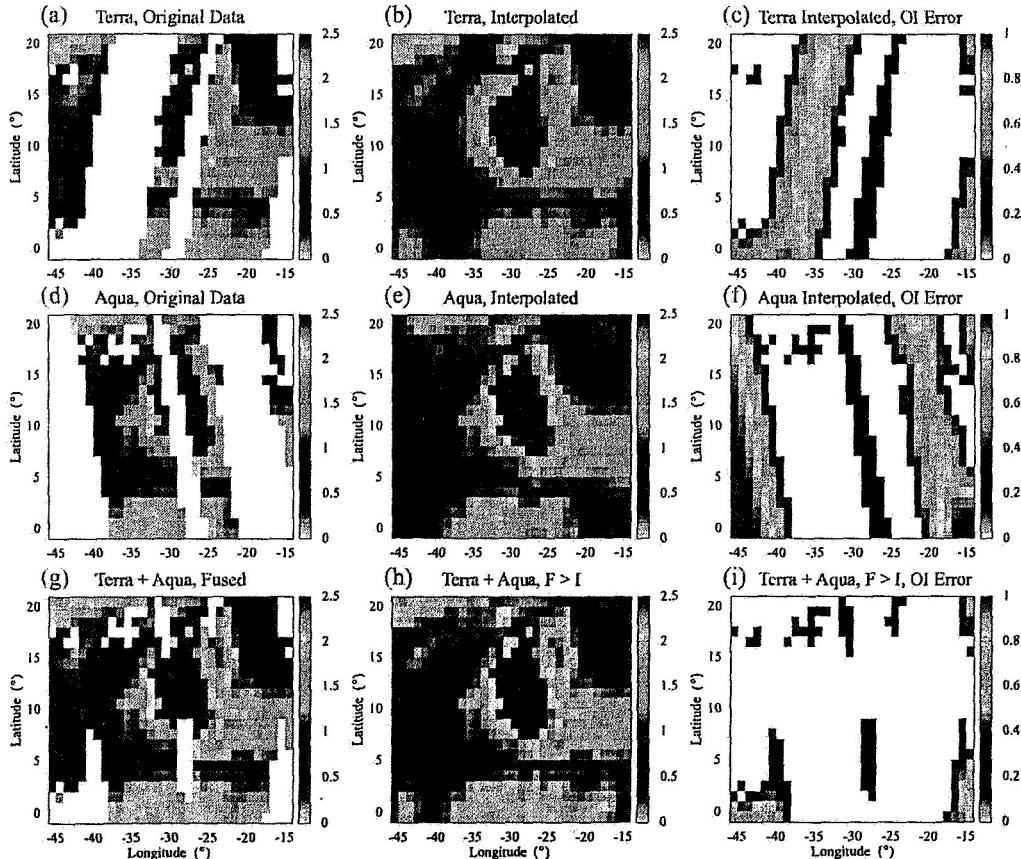


Figure 1. Aerosol Optical Depth on March 14, 2006 for an area with mostly regular gaps (Subset 1): (a) and (d) are the original Terra and Aqua data, respectively; (b) and (e) [(c) and (f)] are the results for Terra [Aqua] interpolated in the gaps separately and the relative errors due to Optimal Interpolation (OI); (g) is the results for Terra and Aqua merged; (h) and (i) are the results for Terra and Aqua fused and then interpolated in the gaps and the relative error due to OI.

REFERENCES

J. Acker, and G. Leptoukh, "Online analysis enhances use of NASA earth science data," *EOS, Transactions, American Geophysical Union*, vol. 88, p.p. 14-17, 2007.

F.P. Bretherton, R.E. Davies, and C.B. Fandry, "A technique for objective analysis and design of oceanographic experiments applied to MODE-73," *Deep-Sea Research*, vol. 23, p.p. 559-582, 1976

C.C. Chu, and J.K. Aggarwal, "The integration of image segmentation map using region and edge information," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, #12, p.p. 1241-1252, 1993.

N.C. Hsu, S-C.Tsay, M.D. King, and J.R. Herman, "Aerosol properties over bright-reflecting source regions," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, #3, p.p. 557-569, 2004

L.A. Remer, Y.J. Kaufman, D. Tanre, S. Mattoo, D.A. Chu, J.V. Martins, R.-R. Li, C. Ichoku, R.C. Levy, R.G. Kleidman, T.F. Eck, E. Vermote, and B.N. Holben, "The MODIS aerosol algorithm, products and validation," *Journal of the Atmospheric Sciences*, vol. 62, p.p. 947-973, 2005.

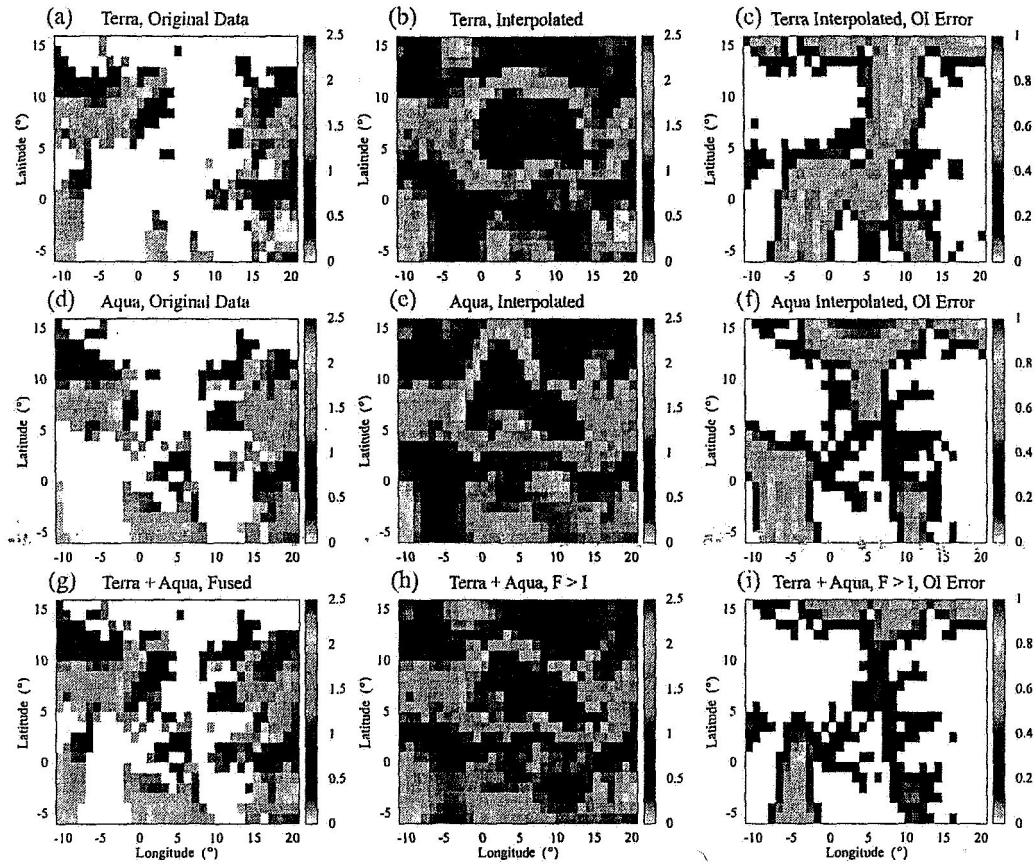


Figure 2. Aerosol Optical Depth on March 14, 2006 for an area with mostly irregular gaps (Subset 2): (a) and (d) are original Terra and Aqua data, respectively; (b) and (e) [(c) and (f)] are the results for Terra [Aqua] interpolated in the gaps separately and the relative errors due to Optimal Interpolation (OI); (g) is the results for Terra and Aqua merged; (h) and (i) are the results for Terra and Aqua fused and then interpolated in the gaps and the relative error due to OI.